ARI Contractor Report 98-01

Selection for Multiple Jobs from a Common Applicant Pool

Cecil D. Johnson
Joseph Zeidner
Dora Scholarios
George Washington University

19980210 099

This report is published to meet legal and contractual requirements and may not meet ARI's scientific or professional standards for publication.

January 1998

United States Army Research Institute for the Behavioral and Social Sciences

Approved for public release; distribution is unlimited

DTIC QUALITY INSPECTED 3

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy) January 1998	2. REPORT TYPE Final Report	3. DATES COVERED (from. . . to) September 1993 - July 1996
4. TITLE AND SUBTITLE Selection for Multiple Jobs from a Common Applicant Pool		5a. CONTRACT OR GRANT NUMBER MDA903-93-K-0020 5b. PROGRAM ELEMENT NUMBER 611102
6. AUTHOR(S) Cecil D. Johnson, Joseph Zeidner, Dora Scholarios		5c. PROJECT NUMBER B74F 5d. TASK NUMBER 1901 5e. WORK UNIT NUMBER C23
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) The George Washington University Office of sponsored Research 2121 I St., NW, Ste. 601 Washington, DC 20052		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 5001 Eisenhower Avenue Alexandria, VA 22333-5600		10. MONITOR ACRONYM ARI 11. MONITOR REPORT NUMBER CR 98-01

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES Contracting Officer's Representative: Peter J. Legree. This report is published to meet legal and contractual requirements and may not meet ARI's scientific or professional standards for publication.

14. ABSTRACT *(Maximum 200 words)*: Procedures for selecting recruits from a common applicant pool to make assignments to a set of 9 or 14 MOS (as surrogates of job families) are evaluated in an unbiased simulation design. Synthetic test scores are generated based on Project A data. Five levels of an assignment strategy level ranging in complexity from one in which jobs and individuals are considered in random order to one which approaches an LP algorithm in both complexity and efficiency. A sixth level is also considered, a primal LP algorithm, as a base-line against which to compare mean predicted performance (MPP) scores provided by the other multiple job assignment procedures. Least squares estimates (LSEs) of the criterion, separately for all 6 strategy facet levels, use 28 Project A tests as predictors. LSEs are used as assignment variables when the "best" weights are obtained from a back sample and as evaluation variables from which to compute MPP when the weights are obtained from the designated population. Two types (levels) of minimum cut scores, one closely resembling the Army operational cut scores with regard to range and height, and the other proportional to dual parameters, are used in conjunction with the 6 levels of the strategy facet. Two sets of assignment variables (AVs), with and without the effect of Brogden's removed, are compared. AVs based on LSEs are also compared with AVs derived from three different types of a single factor. A consistent increase in MPP is found as the complexity of the multiple job selection algorithms approaches the complexity of the LP algorithm. Cut scores proportional to dual parameters provide higher MPPs than the "operational"

15. SUBJECT TERMS

Model Sampling Experiments, Army job families, MPP, Project A, ASVAB, MOS, Personnel selection and classification, SQT Classification and Assignment, Differential Assignment Theory

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unclassified	54	Michael G. Rumsey, 617-275

14. ABSTRACT Continues:

cut scores. Results also show that AVs orthogonal to Brogden's g have greatly reduced intercorrelations while not being reduced in MPP value in the sample in which assignments are made. A potential advantage in using dual parameters obtained from independent samples in conjunction with a LP algorithm is demonstrated.

FOREWORD

This report describes research on the design of efficient personnel systems for effecting the classification and assignment of recruits to MOS. Both the future personnel systems envisaged in this effort and the research approach itself are based on differential assignment theory (DAT). This research adds to the experimentally validated knowledge of DAT principles.

The objective of this research is to provide answers to several critical questions bearing on the classification efficiency of a system, including: (1) methods for reducing the intercorrelations among test composites at a minimum cost to classification efficiency; (2) ways of identifying and reducing shrinkage of classification efficiency that is the consequence of capitalizing on sampling error in the optimal assignment process; (3) the cost in terms of reduced classification efficiency of using multiple job selection procedures instead of a mathematically optimal assignment algorithm, and; (4) the feasibility of using parameters obtained from independent ("back") samples to obtain near optimal personnel assignments, one individual at a time.

SELECTION FOR MULTIPLE JOBS FROM A COMMON APPLICANT POOL

EXECUTIVE SUMMARY

Objectives

This report describes research in the context of differential assignment theory on ways to improve the classification efficiency of personnel systems for effecting the classification and assignment of recruits. This research is conducted in the context of differential assignment theory (DAT) and adds to the experimentally validated knowledge of DAT principles.

The primary goals of this study are as follows:

1. To determine the cost in classification efficiency of using assignment methods similar to traditional selection, and to add complexity and mathematical rigor to this basic process until the linear programming (LP) algorithm is reached;
2. To determine the effect of analysis sample size on the shrinkage of classification efficiency due to the capitalization on error in computing "best" weights and other parameters;
3. To demonstrate that removing the effect of Brogden's *g* from assignment variables (AVs) prior to optimizing assignments on the basis of AVs, considerably reduces the magnitude of intercorrelations among AVs;
4. To explore the feasibility of using traditional weight stabilization techniques (that have been shown to increase cross validation validities) as a means of also increasing population estimates of classification efficiency.

Method

The research approach for all three experiments includes the use of Project A data, using 28 experimental tests as predictors of a core technical proficiency criterion that measures MOS specific knowledge and hands-on performance. A set of 9 MOS are used as surrogates of the Army's operational job families. Five additional MOS (job families) add heterogeneity to this set, but are measured only in terms of job knowledge.

The empirical data are used to compute both validities, corrected for restriction in range, and the intercorrelations among the predictors. These parameters identify the designated population and are used to compute the test weights that are applied to cross sample test scores to provide the evaluation variables (EV). These EV scores provide a population estimate of predicted performance. Classification

efficiency is defined as the sum of the EV scores, converted to statistical standard scores, after assignment of individuals to MOS. This index of classification efficiency, the mean EV, is referred to as mean predicted performance (MPP).

The designated population parameters are also used to compute synthetic test scores to comprise four separate analysis samples and 20 independent cross samples. Weights, parameters and decisions are made in the former; both assignments are made and MPP computed in the latter. Weights computed in the analysis samples are applied to the total set of 28 predictor test scores in the cross samples to form test composites. These composites provide the assignment variables (AVs) for the assignment algorithms utilized in the cross samples. As the first step in computing MPP, population weights are then applied to the same test scores.

A separate set of assignment simulations are conducted for each condition defined as a combination of facet levels. Phase 1 of the first experiment has 36 conditions defined in terms of 6 levels of the multiple job selection strategy facet, 3 levels of a sample size facet, and 2 levels of a minimum cut score facet.

In Phase 2 of the first experiment three types of AVs defined as the product of univariate factor scores and the validities of MOS against the criterion are compared with each other and with AVs consisting of least square estimates (LSEs) of each MOS criterion variable. In additional subsidiary experiments several types of AVs modified to stabilize regression weights or reduce intercorrelations among AVs are compared against LSE type AVs.

Findings

Each step in increased complexity and mathematical precision represented by the 6 levels of the multiple job selection facet provides a further gain in classification efficiency as measured by MPP. The reduction in MPP from using an algorithm that provides separate one at a time assignment of individuals, as contrasted with a primal LP algorithm making Batch Assignments is not large enough to discourage the use of the former. The preliminary application of minimum cut scores based on dual parameters from an independent source increased MPP over the use of the LP algorithm without the use of cut scores.

LSE type AVs that have had the effect of Brogden's *g* removed in the assignment samples provide the same MPP as do unmodified AVs, but the intercorrelations among the AVS are considerably

reduced with the removal of the effect of Brogden's g. The MPPs used in this comparison are unbiased in that MPP is computed from Evaluation Variables (EVs), rather than from the AVs that are utilized by the LP algorithm in making the optimal assignments.

SELECTION FOR MULTIPLE JOBS FROM A COMMON APPLICANT POOL

Table of Contents

Introduction 1
- Need for Research 1
- Background 3
- Research Objectives 6

Experiment 1 7
- Research Design 7
  - The Model Sampling Paradigm 7
  - Project A Predictors, Criterion and Job Sample 11
- Procedure 12
  - Generation of Population Parameters and Analysis Samples 12
  - Estimation of Dual Parameters 13
  - Selection and Assignment Simulations 14
  - Evaluation of Experimental Conditions 18
- Results 19
  - Mean Predicted Performance Standard Scores (MPP) 19
  - Statistical Tests 22

Experiment 2 24
- Research Design 24
- Results 26

Experiment 3 28
- Background 28
- Method 31
  - Research Design 31
- Results 32

Discussion and Conclusions 33
- Unidimensional Assignment Variables 34
- Multiple Job Selection Strategies 35
- Minimum Cut Scores 37
- Four Kinds of Unidimensional Factor Scores 40
- Assignment Variables Orthogonal to Brogden's *g* 41
- Conclusions 43
- Recommendations 44

References 45

List of Tables

Table 1. *Experimental design (Experiment 1)* 8

Table 2. *Project A longitudinal validity predictor measures* 11

Table 3. *Military Occupational Specialties, sample size & operational cut scores* 12

Table 4. *Phase 1 Mean Predicted Performance (MPP) standard scores from classification for 20 replications* 20

Table 5. *Phase 2 Mean Predicted Performance (MPP) standard scores from classification across 20 replications* 22

Table 6. *Combined independent tests of significance for assignment strategies in Phase 1* . . 23

Table 7. *Combined independent tests of significance for differences between operational & proportional cut scores under different assignment conditions in Phase 1* 24

Table 8. *Experimental design (Experiment 2)* 26

Table 9. *Experiment 2 simulation results across 20 replications* 28

Introduction

NEED FOR RESEARCH

The tests of the Armed Services Vocational Aptitude Battery (ASVAB) have been firmly established as valid predictors of Army job performance. A large body of literature that links test scores to the performance of important job tasks is available for forming test composites associated with specific MOS or job families. The effect on mean predicted performance of using these composites to make a first stage selection, from among the applicants into the Army, can be readily calculated using analytical models. A second stage of the Army's personnel selection classification and assignment system to effect the initial assignment of new soldiers to over 250 different entry-level MOS located at over 60 different geographic locations is more difficult to either optimize or assess. Research to design optimal classification and assignment systems, and to measure the resulting effects in terms of a change in average soldier performance, requires the simulation of each alternative system under consideration. Classification efficiency, measured in terms of mean predicted performance, cannot be measured without conducting such a simulation.

Traditionally, the selection of recruits into the Armed Forces is accomplished by establishing minimum requirements (including separate minimum AFQT scores for graduates and non-graduates of high school. Once accepted by a service, a separate set of minimum requirements, including minimum cut scores for ASVAB test composites, is established for initial classification to each MOS. Since raising minimum cut scores on test composites for one MOS reduces the number of higher scoring new soldiers available for other MOS, selection for multiple jobs from a common pool of applicants is very different than selecting from independent pools of applicants available for each MOS. Thus system effects of multiple job selection has more similarity to optimal assignment models (where each person's best composite score has more effect on the classification decision than does comparing the person's score with other individuals) than to stage one selection. Multiple job selection cannot be properly evaluated using a predictive validity index, but instead requires that research be conducted in the context of differential assignment theory (DAT).

DAT is defined by a set of basic concepts and assumptions and further characterized by a set of principles. The major concepts of DAT include the use of mean predicted performance as the measure of personnel system efficiency and the use of completely unbiased research designs that go beyond the

traditional cross validity designs usually considered adequate for the evaluation of selection systems. DAT principles provide a basis for the design and development of effective operational selection and classification systems and for the conduct of research to extend and evaluate DAT.

The more important DAT principles applying to system design include the substitutability of predicted performance for performance scores in many situations important to research on personnel classification, the desirability of having assignment variables (AVs) defined as least squares estimates of the criterion using all available predictors, full least squares estimates (FLSEs), having as many predictors and job families as the data permit and selecting predictors with the maximum amount of differential validity. The existing state of knowledge regarding DAT, as described in Johnson and Zeidner (1995), focuses on the design of a personnel classification system that has well established optimal features such as the use of FLSEs in the context of an optimal assignment model. It is important to extend these principles, with appropriate modifications, to multiple job selection from a common applicant pool, and to the ways that patterns and mean heights of minimum cut scores (viewed in the context of multiple job selection procedures) can be optimized to increase the MPP provided by alternative personnel systems.

Research designs have been proposed in prior DAT related investigations for reducing the negative contribution of sampling error by: (1) specifying LSE-based AVs; (2) optimal clustering of MOS into job families (Johnson, Zeidner, & Leaman, 1992); and (3) selecting best tests from an experimental battery to form an operational classification battery and subsequent test composites (Johnson, Zeidner, & Scholarios, 1995). There is a need for an extension of DAT to provide multiple job selection and LP algorithms and demonstrations of research designs, that have proven potential for increasing the stability of optimal assignments through the use of dual parameters computed in independent (analysis) samples.

The purpose of this study is to identify the most promising directions for further research on DAT rather than to identify specific features for future operational systems. Such research focuses on concepts and potential system changes where further, more applied, research can provide gains in MPP, while providing increased usefulness of the test composite scores for such collateral functions as job counseling, and also providing a safety net to avoid assigning personnel to critical MOS who are highly likely to fail.

The authors would not suggest the use of a primitive multiple job selection system for operational use although one primitive system evaluated in the present study closely resembles a system that has been in operational use. The minimum cut score system presently used by the Army implies the use of the most primitive of these multiple job selection systems. We believe, however, that it is important for managers to be aware of the cost in reduced classification efficiency that results from the use of minimum cut scores.

Personnel system features addressed in this study that the authors believe will eventually be included in future classification systems include the use of parameters from independent samples to effect optimal assignments and the use of assignment variables that are orthogonal to Brogden's *g*. The former of these innovations improves the feasibility of using one-line-at-a-time LP algorithms permitting the immediate separate optimal assignment of each person as the recruit negotiates with a classification counselor. The latter innovation provides a promising way to provide a gain in classification efficiency (MPP) and improves the credibility of the AVs when used as a basis for vocational counseling.

BACKGROUND

Previous studies conducted by the authors examined a variety of selection and assignment conditions under the limiting conditions of using LSEs as assignment variables in a process of optimal assignment to jobs. The focus on personnel classification in terms of a linear programming (LP) type optimal assignment model provides little insight on a multiple job selection models. There remains a large class of unresolved issues bearing on the best way to apply a multiple job selection model to accomplish personnel distribution in an operational system. The present study addresses issues of multiple job selection models.

A multiple job selection model is one in which separate test composites are used to select for two or more jobs with a minimum cut score stipulated for each job. Each time an applicant is compared to a particular job either selection or rejection occurs. This process can be made more efficient by prescribing the order in which either jobs or applicants, or both, are involved in such a comparison. The composites corresponding to each job may be either unique to that job or represent a family of jobs. The minimum cut score will under most systems differ for the various jobs within a job family. Test composites may consist of unit weighted, best weighted, or between unit weighted and best weighted to various degrees of approximation. Cut scores can also be denoted on adjusted scores obtained by subtracting each applicants highest test score from each of his other test scores.

Multiple job selection cannot be expected to closely approximate the amount of utility, measured by MPP, as can be provided by LP algorithms used to optimally assign selected personnel to jobs, (maximizing MPP), while meeting job quotas. However, various multiple job selection models have been utilized as being more feasible to implement, for various policy and administrative reasons, than a more efficient LP based system. One of the most familiar of multiple job selection processes may well be one in which no use is made of the individual test composites taken from a classification battery other than to apply minimum cut scores to the appropriate aptitude area test composite corresponding to each job. A cut score for a given job is usually determined in terms of the judged "difficulty" of that job. This difficulty may reflect the estimated failure rate in school or on-the-job. But to varying degrees this difficulty will usually also reflect the perceived importance of the job to the unit's mission or the safety of it's members. For example, a helicopter mechanic may be given a higher cut score than a tank mechanic because the crew has further to drop in case the engine fails.

A second issue explored in the present study relates to the effect of partialling out Brogden's g from assignment variables and determining the effect on classification efficiency. Brogden (1959) proposed a model which assumed that a component variable which we denote as Brogden's g makes no contribution to classification efficiency. It is possible to define and compute a test score component of most sets of predictor variables which has the properties of Brogden's g. It would also be possible to compute the percentage of each predictor's variance that can be attributed to Brogden's g. The present study provides a demonstration, via an experimental simulation, showing that extracting Brogden's g from assignment variables does not reduce the effectiveness of a classification model. Brogden, however, developed other proofs showing that the use of an LSE as an assignment variable maximizes the expected value of MPP.

The comparison of an H_d based univariate assignment variable with two other assignment variables (one g based and the other H_a based) are closely related to each other and both maximize predictability (one in test space and the other in joint space). This three way comparison is closely related to the second issue.

The third issue addressed in this study is whether methods for increasing the cross sample stability of regression weights have potential for increasing classification efficiency in cross samples, or conversely, whether all or most of these methods are inconsistent with an effort to obtain a set of assignment variables with high differential validity and low intercorrelations among assignment

variables, The investigators considered a number of possible approaches but did not continue into the simulation mode when a particular approach did not appear promising during earlier analytical stages.

Brogden's *g* is defined as an hypothetical variable, or factor, defined in the joint predictor-criterion space. This variable consists of a weighted composite of the predictor tests which is equally correlated with the criterion (i.e., CTP) score of each of the 9 or 14 jobs (MOS). A traditional measure of "Spearman's" *g* would explain more of both the covariances among the AVs and between the AVs and the criterion, and would differ considerably in validity across MOS.

Brogden's MPP has the advantage that it can be computed analytically without requiring a simulation of optimal assignment, but, in contrast to the value of MPP obtained from a simulation, has the disadvantage of being based on a number of limiting assumptions (Brogden, 1959). The estimates of MPP obtained from simulations in this study are obtained from unbiased designs that do not capitalize on sampling error, while Brogden's MPP is computed using a biased design; his parameters are obtained from the sample for which MPP is being computed. Alley and Darby (1995) extended Brogden's table of MPP from 10 to 500 jobs using a Monte-Carlo approach that tended to confirm Brogden's finding, except the authors conclude that Brogden's values underestimated potential benefits.

Optimal assignments to jobs are accomplished in this study separately by both primal and dual linear programming algorithms. In the primal algorithm, the sum of the AVs corresponding to the job to which each person is assigned (the objective function) is maximized with the constraint that the quotas for each job must be met. In the dual algorithm this constraint becomes an objective function while the sum of predicted performance is a constraint (a constant at every iteration). The optimal regions method described by Dwyer (1954) is such a dual algorithm and the allocation constants described by both Brogden and Dwyer (Brogden, 1946, 1954) are different names for dual constants associated with the dual algorithm (Gass, 1975).

The primal LP algorithm can be readily utilized to compute the allocation constants. In a dual LP solution each person has his/her largest AV score subtracted from each of his AV scores to form a set of adjusted scores. The allocation constants, one for each job, can then be added to the adjusted AV scores to form an assignment score for each job. Each person is assigned to the job corresponding to his/her largest assignment score to provide a dual LP solution. Except for rounding errors and for the occurrence of a few ties between assignment scores, the primal and the dual LP algorithms will result in the same optimal assignments. One can compute the allocation constants by first obtaining the primal

LP solution and using these optimal assignments to separate the personnel by jobs. Persons are then placed in sort on adjusted scores and the adjusted score which permits meeting job quotas is identified. That score is the allocation constant.

Minimum cut scores on the AVs used in multiple job selection models to separately reject applicants for each job have been traditionally associated with the perceived difficulty and/or importance of each job. If the dual parameters (allocation constants) are computed in a back sample and applied in a cross (assignment) sample, the use of cut scores proportional to these dual parameters in a multiple job selection model and the use of the actual allocation constants in a dual LP algorithm provide similar assignments.

RESEARCH OBJECTIVES

The more significant of the research objectives of the present study include:

1. Applying DAT methodology for evaluating classification efficiency of multiple job selection approaches.
2. Obtaining an increased understanding of the impact of analysis (back) sample size on MPP shrinkage, the standard deviations of MPPs across experimental replications, validities and intercorrelations of AVs, and the relationship of Brogden's MPP with MPP measured in a simulation process.
3. Verifying that partialling out Brogden's *g* from a set of AVs does not reduce the classification efficiency of the adjusted AVs.
4. Exploring the effect of partialling out Brogden's *g* from AVs on a number of key characteristics of the classification model, including value of MPP after optimal assignment, values of Brogden's MPP, and SDs of Brogden's MPP across replications.
5. Determining the feasibility of using weight stabilization concepts to improve the effectiveness of sets of AVs in cross samples when the analysis samples are small.
6. Improving understanding of the magnitude and direction of changes in MPP resulting from two alternative approaches to the use of cut scores.

Experiment 1

RESEARCH DESIGN

The Model Sampling Paradigm

The experimental simulations of the Army selection and assignment process were achieved using a model sampling approach for the generation of population and sample data. This approach was used in earlier Army manpower allocation simulations (e.g., Niehl & Sorenson, 1968; Olson, Sorenson, Hayman, Witt, & Abbe, 1969; Sorenson, 1965) and extended recently to provide a rigorous cross validation design framework for simulation experiments evaluating the outcomes of alternative selection and assignment procedures (see Johnson and Zeidner, 1991).

Samples of synthetic entities are generated to have statistical characteristics in common with an empirical sample which has been designated as the applicant population for the simulation. In the manner described by Johnson and Zeidner (1991, p. 127), this approach allows the generation of any desired number of random samples representative of the "designated" population.

Either separate samples of synthetic entities or independent sources of empirical score vectors are used for the following purposes: (1) computing parameters and or selecting variables to be used in simulating personnel classification and assignment; (2) the conduct of the selection and/or classification simulations; and, (3) evaluation of the utility of the alternative procedures in terms of mean predicted performance (MPP). The samples, either obtained from empirical data or consisting of generated entities, used for these separate purposes are referred to, respectively as analysis, cross validation or evaluation samples. The experiment consists of a simulation conducted in the cross validation samples that uses the variables and parameters obtained from the analysis sample and subsequently evaluated in the cross samples using weights obtained from either an analysis sample or the designated population.

The simulations conducted in the cross samples of this study are replicated a number of times to provide further stability. Twenty replication samples of entities are utilized in all the experiments reported here. Experimental conditions are evaluated by the mean predicted performance of the assigned samples, computed using weights based on the population parameters. The same evaluation weights are utilized across all experimental conditions. The classification process simulated in these experiments includes 4 levels of multiple job selection and 2 levels of optimal assignment to jobs by means of a linear programming (LP) algorithm. In previous studies all assignments were simulated using an LP algorithm.

The model sampling approach, as summarized above, has been used in other recent personnel selection and classification simulation experiments using the Project A data (see for example, Statman, 1992; Scholarios, Johnson & Zeidner, 1994). Corrections for restriction in range are based on the 1980 reference youth population.

Experimental Design

Experiment 1 involved a total of 76 experimental conditions simulating selection and assignment to Project A jobs. These conditions were applied in two phases. The first phase examined variations in: (1) the strategy for assigning individuals to jobs; (2) the type of job cut scores for determining minimally eligible candidates for assignment; and (3) the size of the analysis sample used to calculate test and predicted performance scores for applicant samples. In all the first phase simulations, assignment was based on predicted performance estimated by the job-specific least squares estimate (LSE) score calculated on the full set of Project A predictors. Phase 2 introduced a further experimental variable to enable a comparison of alternative assignment variables against this LSE assignment variable. The four experimental variables in Phases 1 and 2, and their levels, are summarized in Table 1 and described in detail below.

Table 1. *Experimental design (Experiment 1)*

	Phase 1 Variables
Assignment strategy	1. Random order assignment 2. Rank-ordered jobs based on validity 3. Multidimensional hierarchical classification (HC) 4. Cascading assignment 5. Multidimensional selection (MDS) 6. Optimal assignment using a linear programming algorithm
Cut score	1. Operational cut scores 2. Cut scores proportional to MDS column constants
Analysis sample size	1. N = 1620 2. N = 2430 3. N = 3645
	Phase 2 Variable
Assignment variable	1. Least squares estimate (LSE) of criterion 2. First principal component (g) 3. Horst's index of absolute validity (H_a) 4. Horst's index of absolute validity (H_d)

Selection/assignment strategy

Each of the assignment strategies was designed either to consider entities separately for assignment or to consider jobs and/or entities in a prescribed order. All assignment strategies were performed on the samples of minimally eligible entities after selection and consideration of job cut scores. Six strategies were applied to the selected samples.

1. Random ordering assignment considers entities and fills jobs in random order, using random numbers to determine the order in which jobs are to be filled.
2. Rank-ordering based on the predictability of job assignment variables allows jobs to be filled in the order of their validity, making a random selection from the minimally eligible individuals.
3. Multidimensional hierarchical classification (HC) involves the rank ordering of both jobs, on their predictability, and eligible individuals, in terms of their predicted performance. Assignment is based on considering each job in order and filling jobs by top-down selection of individuals drawn in order from the predicted performance score continuum corresponding to that job until the quota is met.
4. Cascading assignment rank orders jobs on the height of their cut score and eligible individuals in terms of their predicted performance corresponding to each job being considered.
5. Multidimensional selection (MDS) assigns individuals on the basis of their highest adjusted predicted performance scores for each job calculated by adding a constant for each job (also known as column constants or dual parameters) to each individual s predicted performance scores. The optimal set of column constants is obtained by implementation of an optimal, primal assignment algorithm to determine the solution which achieves all quotas and ensures that no rejected individual has a higher predicted performance for a particular job than those who are assigned.
6. Optimal linear programming assignment maximizes the mean predicted performance (MPP) of the assigned group.

Cut scores

Prior to implementing the assignment strategies, minimally eligible individuals were identified by comparing predicted job performance scores to the relevant job cut scores. Two experimental sets of cut scores for the Project A jobs were applied to the applicant samples. A third level of this facet is described below for comparison purposes. This third level was seriously considered during the initial stage of Experiment 1 but was deleted and not further implemented.

1. Operational cut scores were converted from Army standard scores to a scale corresponding to statistical standard scores with a mean of zero and standard deviation of one. Thus these converted cut scores became appropriate to apply to the synthetic entities used in the simulations. The range of the converted cut scores across 14 surrogate MOS is .75.

2. The job constants calculated for the purposes of MDS assignment were transformed to have a mean of zero and a standard deviation of one across the set of constants for the 14 jobs selected as surrogates for the operational job families. This change in scale retains the same relationships, rank order wise, across the set of surrogate MOS as possessed by the dual parameters calculated for each MOS in the MDS calculations. The range of the converted cut scores across 14 surrogate MOS is 2.6 and is as large as 3.6 for 9 surrogate MOS based on the smallest analysis sample.

3. The job constants calculated for the purposes of MDS assignment were transformed to have the same mean and standard deviation as the operational cut scores for the set of surrogate MOS. As with level 2, this change in scale retains the same relationship, rank order wise, across the set of surrogate MOS as possessed by the dual parameters calculated for each MOS in the MDS calculations. If MOS quotas were equal, and the means and standard deviations of these cut scores were equal, the same percentage of applicants would be rejected as by the operational cut scores.

Analysis sample size

Analysis samples of predictor intercorrelations and validity coefficients, which under the model sampling paradigm were assumed to be random samples from the applicant youth population, were used to calculate the regression weights required to obtain predicted performance scores for each job. The analysis samples were calculated by aggregating across synthetically-created samples of random normal deviates for each of the jobs represented in the Project A data. Three different sample sizes of entities for each job were used in this aggregation process: Ns of 1620, 2430 and 3645.

Assignment variable

The assignment variable experimental treatment contrasted three types of univariate variables appropriate for use in a hierarchical classification model with the vectors of least squares estimates providing individuals predicted performance scores for each job for use in an allocation model as defined by Johnson and Zeidner (1991, p. 17). This provided four alternative types of predicted performance scores to be used for assignment.

1. The least squares estimate (LSE) composites of the full set of predictors computed separately for each job for use in the allocation model.
2. The first principal components factor which is representative of the traditional measure of general cognitive ability (g) for use in the hierarchical classification model.
3. Horst's index of absolute efficiency (H_a) factor maximized in the joint predictor-criterion space for use in the hierarchical classification model.

4. Horst's index of differential efficiency (H_d) maximized in the joint predictor-criterion space for use in the hierarchical classification model.

DATA

Project A Predictors, Criterion and Job Sample

The predictor measures, criterion and validation data from the longitudinal validation phase of Project A were used to simulate selection and classification to Project A jobs. Table 2 shows the predictor measures used in this experiment. These include the nine tests of the Armed Services Vocational Aptitude Battery (ASVAB) and an additional 16 experimental predictors included in the longitudinal validation phase. As noted earlier, the criterion used for the validation data was the core technical proficiency criterion measure for each job. Jobs were represented by the 14 Military Occupational Specialties (MOS) for which longitudinal validation data was available (see Table 3). The nine Batch A jobs were used in Phase 1 of the experiment while additional conditions in Phase 2 were completed using the full set of 14 jobs as shown in Table 3.

Table 2. *Project A longitudinal validity predictor measures*

ASVAB tests	Experimental tests
1. General Science	10. Basic Accuracy Composite
2. Arithmetic Reasoning	11. Perceptual Speed Composite
3. Numerical Operations	12. Perceptual Accuracy Composite
4. Coding Speed	13. Assembling Objects
5. Auto Shop information	14. Job Composite: Routine
6. Mathematical Knowledge	15. Avoice Composite: Administrative
7. Mechanical Comprehension	16. Avoice composite: Food Service
8. Electronics Information	17. Avoice composite: Structural/Machines
9. Verbal Ability	18. Avoice composite: Protective Services
	19. Avoice composite: Rugged/Outdoors
	20. Avoice composite: Skilled Technical
	21. Factor Score: Locus Control
	22. Factor Score: Cooperate
	23. Factor Score: Dominance
	24. Factor Score: Dependability
	25. Factor Score: Physical Condition
	26. Factor Score: Stress Tolerance
	27. Factor Score: Work Orientation

Table 3. *Military Occupational Specialties, sample size & operational cut scores*

Military Occupational Specialties (MOS)		N	Cut scores[a]
	Batch Z		
12B	Combat Engineer	487	85
16S	MANPADS Crewman	247	90
54B	ST Chemical Operator	275	95
76Y	Unit Supply Specialist	473	90
94B	Food Service Specialist	515	90
	Batch A		
11B	Infantryman	263	85
13B	Cannon Crewmember	607	85
19K	CO M-1 Abrams Armor Crewman	486	85
31C	Single Channel Radio Operator	186	100
63B	Light Wheel Vehicle Mechanic	441	95
71L	Administrative Specialist	273	95
88M	Motor Transport Operator	236	90
91A	Medical Specialist	566	95
95B	Military Police	276	95

[a] In Army standard scores with mean = 100 and standard deviation = 20

PROCEDURE

Generation of Population Parameters and Analysis Samples

Population predictor intercorrelations and validity coefficients were obtained by correcting the Project A empirical data for restriction in range due to selection and classification. The empirical Project A longitudinal validity data for the 27 predictors and 14 jobs were corrected back to the 1980 reference youth population (Mitchell and Hanser, 1980). The youth population ASVAB intercorrelations provided the basis for the correction procedure which used the formulae of Gulliksen (1950, p. 165, numbers 37 and 42) and is described in Johnson and Zeidner (1991). The corrected predictor intercorrelations and validity coefficients, calculated both for the Batch A 9-job set and the combined Batch A and Z 14-job set, were designated as the applicant population for this series of simulations.

Analysis samples, used in the present experiment to compute the job dual parameters and regression weights for assignment variables, were generated to have the same statistical characteristics as the population parameters. The procedure for generating analysis samples followed that described in Johnson, Zeidner and Scholarios (1990) and Statman (1992). In the present experiment, three sets of

analysis samples of different sizes were created to represent different treatments. These were achieved by generating matrices of random normal deviates for each job with dimensions N by $n + 1$, reflecting the alternative sample sizes (N = 1620, N = 2430, or N = 3645) and the number of predictors (n = 27) plus the single core technical proficiency criterion variable. These matrices formed the starting point for the transformation procedure which produced different analysis samples of predictor intercorrelations and validity coefficients for the 27 predictors and 14 Project A jobs.

Estimation of Dual Parameters

Sets of dual parameters (known as column constants by those familiar with the derivations and discussions of the "optimal regions" algorithm first proposed by Brogden and later implemented in greater detail by Dwyer) were generated for use in experimental conditions involving the dual assignment algorithm referred to as multidimensional selection, and for the use of cut scores proportional to the dual parameters. Different sets of dual parameters were generated for each of the three analysis sample conditions described in Table 1. The procedure for calculating each set of column constants was based on a primal linear programming assignment process involving independent samples of entities. The use of column constants as parameters in a dual algorithm can provide an assignment solution in the cross samples which maximizes predicted performance while ensuring that all job quotas are met, without the use of a primal linear programming algorithm to make the simulated operational assignments. The following paragraphs describe how a primal LP solution can be used to obtain the column constants required for a dual LP solution.

For the three analysis samples, selection and assignment was separately simulated for samples of synthetic entities of Ns of 1620, 2430 and 3645 entities, respectively. Random normal deviates were transformed into vectors of synthetic test scores based on analysis sample predictor intercorrelations. Each vector represented an entity which also included an Armed Forces Qualifications Test (AFQT) score. Selection was simulated by truncating the distribution of AFQT scores in a manner corresponding to a selection ratio of .70 and discarding all entities with an AFQT score below the point of truncation. Predicted performance scores were calculated for the remaining selected entities using analysis sample regression weights applied to these test scores.

The selected entities were divided into replication subsamples each of which underwent separate assignment. Each subsample was of a size which would allow equal quotas of entities for the 9 jobs (for Phase 1 simulations) or 14 jobs (for Phase 2 simulations). For analysis sample N = 1620 (after selection

N = 1134), three subsamples of 378 allowed quotas of 42 for each of the 9 jobs; for analysis sample N = 2430 (after selection N = 1692), four subsamples of 423 allowed quotas of 47 for each of the 9 jobs; and for analysis sample N = 3645 (after selection N = 2538), six subsamples of 423 allowed quotas of 47 for each of the 9 jobs. For Phase 2 simulations using 14 jobs, a selection ratio of .6914 was applied to analysis sample N = 3645, leaving N = 2520 after selection, and enabling nine assignment subsamples of 280 entities and quotas of 20 for each of the 14 jobs.

For each replication subsample of each analysis sample, a primal linear programming algorithm was used to optimally assign entities to each of the jobs. Considering jobs in turn, and only the entities assigned to the job being considered, adjusted predicted performance scores were calculated for each entity by subtracting their highest scores from each of their other scores. The lowest adjusted predicted performance score held by an entity assigned to that job, multiplied by -1 to provide a positively-signed constant, was designated as the dual parameter for that job. Repeating this procedure for all replication subsamples, the final estimated column constants for each analysis sample were obtained by averaging the constants obtained for each subsample.

The dual parameters were used to compute adjusted predicted performance scores for assignment under the Multidimensional Selection strategy (i.e., use of dual parameters to make optimal assignments to jobs). It was originally intended that these parameters would be adjusted for use as the experimental cut score treatments by standardizing these dual parameters to provide the same mean and standard deviation as the operational cut scores (see Table 3). However, when the numbers of levels were reduced from 3 to 2, level 2 was used instead of level 3. In level 2 the standardized dual parameters for each job, calculated separately for each of the analysis sample conditions, have a minimum range of 2.6 and a maximum range of 3.6. Since the operational cut scores converted to standard score form have a range of .75 over the 14 jobs, the experimental cut scores have standard deviations several times larger than the operational cut scores.

Selection and Assignment Simulations

The cross validation design of this experiment required the generation of multiple random samples with expected values approximating the intercorrelations and validity coefficients of the population. The general procedure for generating applicant cross samples used in recent personnel selection and classification simulations is described in Johnson and Zeidner (1995). In the present experiment, twenty random applicant samples, each consisting of 450 hypothetical job applicants (or

entities), formed the input to the selection and assignment simulations. Each applicant sample was generated from a matrix of random normal deviates with dimensions N by n, where N was the number of applicants (450) and n was the number of predictors (27). These random normal deviates were then transformed to test scores using a Gramian factor solution of the population intercorrelation matrix. The 20 replication samples of test scores were exposed to each of the 36 conditions of Phase 1 and 40 conditions of Phase 2 which simulated selection, calculation of predicted performance scores using alternative analysis sample weights, comparison with job cut scores, and assignment to jobs. Alternatively stated, the entire experiment was replicated on 20 random samples with each experimental condition based on a total sample of 9,000 entities.

Selection was simulated by truncating the sample at the lower tail of the distribution of applicants' Armed Forces Qualifications Test (AFQT) scores. A selection ratio of .70 was applied leaving 315 job applicants within each cross sample to be considered for assignment after selection.

The effect of the cut score was achieved by comparing the entity predicted performance score to the cut score for each job being considered. A negative constant (in this case -3.0) was subtracted from the predicted performance score of each entity score falling below the minimum cut score for that job. Thus, the predicted performance scores failing to meet the minimum requirements for the corresponding job became negative in value but retained their original rank order within the failed group. If all entity scores failed the cut score comparison, the entity with the score closest to the cut score was assigned.

The two levels of the cut score facet were applied to the truncated sample of 315 entities remaining after initial selection. Under the first level of this facet, the operational cut score conditions used the Project A job cut scores (see Table 3) converted from Army standard scores to statistical standard scores. The second level used cut scores made proportional to the dual parameters by standardizing separately across the 9 and 14 sets of MOS to a mean of zero and a standard deviation of one. Note that the operational cut scores, although applied to AV scores having a mean of zero and a SD of one had a mean considerably higher than zero and a SD considerably less than one.

Finally, the assignment simulation was completed using the entity predicted performance scores as assignment variables within each of the six experimental assignment conditions. Assignment variables in Phase 1 of the experiment were calculated as least squares estimates (LSEs) of performance for the 9 Batch A jobs using the full set of 27 predictors. LSEs for each job were calculated for the 315

selected individuals and, after consideration of the minimum requirements of each job cut score, used as the assignment variables in the six different assignment strategies of this experiment.

In Phase 2, three univariate assignment variables were calculated for comparison with the LSE predicted performance scores and used for assignment to the 14 Batch A and Z jobs. The LSE assignment variable was calculated in the same manner as Phase 1 except using the assignment target of the 14 jobs. The three univariate assignment variables were given by the factor composites of the 27 predictor variables weighted by the composites intercorrelations with the 14 job criterion measures (i.e., the factor validity coefficients). Three different factor solutions provided the three univariate assignment variables: (1) a general cognitive ability (*g*) assignment variable was created from a principal components factor solution of the predictor space (2) an additional *g* variable was created by maximizing Horst's index of absolute efficiency (H_a) in the joint predictor-criterion space; and (3) a differential efficiency assignment variable was created by maximizing Horst's index of differential efficiency (H_d) in the joint predictor-criterion space.

Each assignment strategy varied in terms of the priorities given to entities and jobs in the assignment algorithm. These were implemented in the following manner.

1. Random ordering assignment

Random order assignment used a rectangular distribution of random numbers to place jobs in order for assignment. Given that the entities representing individuals were already randomly generated, each entity was assigned in turn (i.e., random order) to the job for which their predicted performance was highest, considering a different random ordering of jobs for each replication and condition of the experiment. Initially jobs were filled only from the group of entities which had met the minimum requirements of each job. If the goal of equal quotas was not met, under-filled jobs were filled randomly from the remaining underqualified entities.

The objective of the assignment algorithm was to create an assignment matrix of zeros and ones with dimensions N by m, where N is the number of selected and assigned entities in each cross sample (N = 315) and m is the number of jobs ($m = 9$ in Phase 1 and $m = 14$ in Phase 2). This matrix represented the person-job match, where 1 signified an assigned entity, for the particular assignment strategy.

2. Rank-ordering of jobs by validity

The second assignment algorithm implemented random assignment of entities as in the first algorithm, differing only in terms of the rank-ordering given to the jobs to be assigned. Jobs were considered for assignment in order of their predictability, which was measured by the diagonal of the covariance matrix computed from the analysis sample predictor intercorrelations and validities. As before, under-filled quotas were filled from entities failing the cut score comparison and an assignment matrix of zeros and ones was created to represent the person-job match.

3. Multidimensional hierarchial classification (HC)

The algorithm for multidimensional hierarchical classification (HC) utilized the same ordering of jobs by their predictability as the second algorithm, but in addition, effected a rank ordering of entities according to their predicted performance scores. The initial assignment procedure considered jobs as indicated by the rank-ordering of their validities and filled the quota for each job in turn, from the top down of the ranked distribution of entity predicted performance scores. Only those entities considered minimally eligible for each of the jobs was included in this initial assignment. Under-filled quotas, caused by high cut scores, were filled using a secondary assignment algorithm, in which unassigned entities were assigned to jobs in the same rank ordering as before (i.e., filling first the job with the highest predictability) and in the rank order of the entity predicted performance scores. Jobs remaining under-filled at the conclusion of the secondary assignment procedure were filled randomly from the remaining group of entities which had not met any of the minimum entry requirements for this set of jobs. As in the other assignment algorithms, an assignment matrix of zeros and ones represented the person-job match resulting from this procedure.

4. Cascading assignment

Cascading assignment replicated the algorithm of multidimensional HC described above, except for the ordering in which jobs were considered. Predictability was measured by the rank-ordering of the magnitude of either the operational cut scores or the dual parameter cut scores, depending on the experimental condition, for each of the jobs being considered. Thus, once again, both jobs and entities were considered in a prescribed rank order, under-filled quotas were filled using the same secondary and random assignment conditions as in hierarchical classification, and an assignment matrix of zeros and ones was produced to represent the person-job match.

5. Multidimensional selection (MDS)

The MDS or dual assignment algorithm followed the procedure described in Johnson and Zeidner (1990, p.48). The dual parameters, obtained from the optimal assignment process described above, were added to each of the predicted performance scores to provide adjusted assignment variables. Individuals were then assigned to the jobs corresponding to their highest adjusted assignment score.

6. Optimal linear programming assignment

A primal LP algorithm (a network optimization model) was used to make an optimal assignment that maximized MPP in each cross sample. As with other assignment conditions, equal quotas were met for each job being considered.

Evaluation of Experimental Conditions

A comparison of the selection and assignment experimental conditions was based on the mean predicted performance (MPP) of the assigned group of entities. This indicator of efficiency was based on the person-job matches resulting in the assignment matrices from each condition and computed using evaluation predicted performance scores rather than the assignment variable scores that were considered in making assignments under each condition. The evaluation scores were calculated using regression weights derived from the designated population intercorrelation and validity coefficients. The predicted performance scores resulting from the relevant person-job assignments for each condition were averaged across the 20 replication samples to produce a MPP standard score for each experimental condition.

The MPP produced from the simulations was an indicator of the total efficiency gained from selection and assignment under these conditions. The efficiency due to classification effects was ascertained by subtracting out the gain in MPP attributable to selection. This was computed as a function of the selection ratio and the validity of the selection variable (in this case the AFQT) in the manner recommended by Naylor and Shine (1965). Using a selection ratio of .70 and an average AFQT validity for 9 MOS in the longitudinal validity data of .4842, yielded an expected MPP attributable to selection of.2372 for Phase 1 simulations. For the Phase 2 simulations using 14 jobs, an effective selection ratio of .69 (given the constraint of equal quotas for 14 jobs) and an average AFQT validity for the 14 jobs in the longitudinal validity data of .5167, the MPP attributable to selection was .8325. This was subtracted from the total MPP gain to provide the MPP attributable to gains in classification efficiency alone.

RESULTS

Mean Predicted Performance Standard Scores (MPP)

Table 4 shows MPPs for three analysis sample sizes. MPP due to classification effects in Phase 1 ranged from .2562 for the use of random order assignment in conjunction with the use of operational cut scores applied to LSE composites to .7762 for the use of optimal assignment and cut scores proportional to column constants. The relatively high magnitude of these MPPs compared to most other model sampling studies based on assignment to only 9 jobs reflects the use of several conditions favorable to a higher MPP as a result of optimal classification. These include the use of LSEs based on 27 predictors for the assignment variables, the use of generated (synthetic) scores, the use of an SR of .30 (instead of .20 or .25 that has been used in some other studies), and the correction for restriction in range with the youth population designated as the target population for the correction. Note that all of these conditions increases MPP due to classification, including the increased selection due to the use of a higher AFQT score to reject more simulated applicants for enlistment.

The 6 levels of the strategy facet were listed in order of their expected classification efficiency and the MPP results were in agreement with this expected hierarchy. The largest increase in MPP occurs between the second and third levels and the smallest difference between the first and second levels. The remaining differences between the adjacent members of this hypothesized continuum are large enough to be both statistically significant and non-trivial.

Table 4. *Phase 1 Mean Predicted Performance (MPP) standard scores from classification for 20 replications*

	Assignment Strategy					
	Random order assignment	Rank ordered jobs by validity	Multi-dimensional HC	Cascading assignment	MDS	Optimal assignment
	Analysis Sample N = 1620					
Operational cut score	.2562 (.040)	.2607 (.023)	.4678 (.029)	.5615 (.035)	.6617 (.040)	.7022 (.038)
Cut scores proportional to column constants	.2745 (.033)	.2868 (.038)	.4732 (.033)	.5691 (.035)	.7129 (.035)	.7497 (.036)
	Analysis Sample N = 2430					
Operational cut scores	.2662 (.023)	.2819 (.026)	.5074 (.039)	.5998 (.035)	.6773 (.042)	.7531 (.038)
Cut scores proportional to column constants	.2945 (.038)	.3052 (.027)	.5077 (.039)	.6014 (.035)	.7019 (.041)	.7717 (.035)
	Analysis Sample N = 3645					
Operational cut scores	.2858 (.026)	.2902 (.028)	.5181 (.032)	.6042 (.037)	.7030 (.037)	.7585 (.037)
Cut scores proportional to column constants	.2867 (.038)	.3034 (.037)	.5198 (.033)	.6043 (.037)	.7096 (.036)	.7762 (.034)

The Phase 2 results, using 14 instead of 9 jobs shown in Table 5 confirmed the general progressive improvement in MPP from each change in the assignment strategy when LSE assignment variables were employed, as observed in Phase 1. However, Phase 2 of Experiment 1 was primarily conducted to examine the effect of an increased number of variables and to inject the alternative use of three kinds of univariate assignment variables.

Comparing Table 4 with Table 5 we see that increasing the number of jobs from 9 to 14 increases MPP from .7762 to .9389 when cut scores proportional to column constants are utilized in

conjunction with optimal assignment and LSEs. There is a similar large increase for using 14 jobs when proportional job constants are used with MDS, and a smaller but significant increase for increasing the number of jobs when Multidimensional HC is used. There is a much smaller reduction in MPP from increasing the number of jobs when the strategy for level 2 is used in conjunction with the use of either type of cut scores.

No entries for random order assignment were computed for Table 5 because unidimensional HC provides a trivial theoretical gain in MPP from an increase in number of jobs, and the "random order" algorithm should be similar in this respect. For a similar reason, no provision was made for comparing the univariate assignment variables (g, H_a, and H_d) as a means of detecting the effect of increasing the number of jobs from 9 to 14.

The most prominent difference in the Phase 2 set of results shown in Table 5 is that the least squares estimate (LSE) assignment variables, as used in the first phase of simulations, were vastly superior to all the univariate assignment variables which had been generated from factor solutions. All comparisons between LSE and factor score assignment variables were significant at $p < .0005$ using the critical ratio for correlated differences and the normal curve table to calculate significance.

It is not surprising that no distinction could be made between the principal component assignment variable (g) and the factor solution in which H_a is maximized, with differences averaging around .002 standard scores. Comparing both these variables to the variable based on a factor solution in which H_d is maximized, only under optimal assignment was there a significant increase in MPP standard score from using H_d. Using the operational cut scores, H_d assignment variables were significantly better than H_a assignment variables ($p < .05$). Using the cut scores proportional to column constants, the H_d based assignment variable provided a higher MPP than the g based assignment variable by .04 MDs ($p < .10$).

Table 5. *Phase 2 Mean Predicted Performance (MPP) standard scores from classification across 20 replications*

Assignment variable	Rank ordered jobs by validity	Multidimensional HC	Cascading assignment	MDS	Optimal assignment
			Operational cut scores		
LSEs	.2886 (.025)	.6197 (.031)	.6177 (.036)	.7480 (.035)	.8262 (.038)
g	.0583 (.044)	.0657 (.046)	.0531 (.036)	.0437 (.047)	.0681 (.052)
H_a	.0628 (.051)	.0612 (.035)	.0548 (.042)	.0751 (.051)	.0467 (.034)
H_d	.0796 (.034)	.0897 (.048)	.0588 (.035)	.0746 (.051)	.0916 (.051)
			Cut scores proportional to column constants		
LSEs	.2840 (.031)	.6965 (.031)	.6201 (.034)	.8394 (.038)	.9389 (.035)
g	.0644 (.049)	.0671 (.045)	.0602 (.037)	.0732 (.039)	.0561 (.042)
H_a	.0614 (.040)	.0622 (.039)	.0561 (.038)	.0682 (.041)	.0665 (.049)
H_d	.0858 (.034)	.0632 (.027)	.0575 (.032)	.0681 (.041)	.0933 (.047)

Statistical Tests

Comparisons between adjacent pairs of experimental conditions are shown in Table 6. These statistical tests use the 20 repeated measure mean predicted performance (MPP) standard scores obtained from classification effects (see Tables 5 and 6). The primary comparisons between MPP for the Phase 1 of Experiment 1 involved tests of significance for: (1) adjacent levels of assignment strategies; and (2) for comparisons between the two cut score methods used under different assignment conditions.

These comparisons were obtained using Fisher's (1948) method for combining independent tests of significance in which critical ratios between correlated MPPs across the 20 replications are first converted to significance levels and then into chi-square values with 2 degrees of freedom. Fisher's method enables the combined statistical significance to be obtained from the summation of both independent chi-square values and degrees of freedom attached to each independent estimate.

The mean differences (MDs) between adjacent levels in the assignment strategy facet result in the differences between the first two levels, random assignment and rank ordered jobs, showing the smallest differences. In conjunction with the use of operational cut scores a non-significant difference of .008 resulted, while the difference of .013 standard scores that resulted from the use of proportional cut scores in conjunction with these two same levels is significant at the .05 level. Beyond this first

comparison, there is a consistent increase in the MPP gained from each incremental improvement to the assignment strategy, as reflected in the levels of this facet. The differences between all levels beyond the first and second levels are significant at beyond $p<.005$, with the largest MDs resulting from comparing the second and third levels and the smallest from comparing the fifth and sixth levels. The facet 1 MDs significant at $p < .005$ ranged between .058 and .220.

Table 6. *Combined independent tests of significance for assignment strategies in Phase 1*

Comparison[a]	MD[b]	CR1[c] (Analysis sample N=1620)	CR2[c] (Analysis sample N=2430)	CR3[c] (Analysis sample N=3645)	X^2(CR1) (df=2)	X^2(CR2) (df=2)	X^2(CR3) (df=2)	combined X^2 (df=6)
			Operational cut scores					
Random v Rank ordered	.008	.41	-1.34	.46	1.83	.21	2.41	4.45
Rank ordered v MHC	.220	24.43	20.79	22.95	>15.2	>15.2	>15.2	>45.6***
MHC v Cascading ass.	.091	9.20	7.50	7.24	>15.2	>15.2	>15.2	>45.6***
Cascading ass. v MDS	.093	8.53	5.81	8.26	>15.2	>15.2	>15.2	>45.6***
MDS v Optimal	.058	3.33	5.94	4.59	15.2	15.2	15.2	45.6
			Cut scores proportional to column constants					
Random v Rank ordered	.013	1.07	1.12	1.43	4.61	3.79	3.79	12.193*
Rank ordered v MHC	.202	19.05	18.42	18.99	>15.2	>15.2	>15.2	>45.6***
MHC v Cascading ass.	.092	5.27	7.50	7.24	>15.2	>15.2	>15.2	>45.6***
Cascading ass. v MDS	.117	12.65	8.02	8.81	>15.2	>15.2	>15.2	>45.6***
MDS v Optimal	.058	3.10	5.51	5.83	13.8	15.2	15.2	44.22***

* $p < .05$ *** $p < .001$

Notes.

[a] **Assignment strategy abbreviations**: Random: random order assignment; Rank ordered: rank ordered jobs by validity; MHC: multidimension hierarchical classification; Cascading ass: cascading assignment; MDS: multidimensional selection; Optimal: optimal linear assignment

[b] **MD** denotes mean difference between MPP standard scores for adjacent assignment strategies (see Table 4) across 3 analysis sample conditions

[c] **CR** denotes critical ratio

The facet 2 comparisons, between the two types of cut score variables (facet 2: operational cut scores and cut scores proportional to the column constants) using Fisher's method, resulted in fewer

significant differences than resulted from the facet 1 comparisons. Table 7 shows that only comparisons between levels 2 and 3 and between 4 and 5 provided significance at p<.01. a third difference, between facets 4 and 5, provided significance at p < .05. The MDs for the three significant facet 2 comparisons were all between .02 and .03.

Table 7. *Combined independent tests of significance for differences between operational & proportional cut scores under different assignment conditions in Phase 1*

Assignment Condition[a]	MD[b]	CR1[c] (Analysis sample N=1620)	CR2[c] (Analysis sample N=2430)	CR3[c] (Analysis sample N=3645)	X^2(CR1) (df=2)	X^2(CR2) (df=2)	X^2(CR3) (df=2)	combined X^2 (df=6)
Random ass.	.020	-.54	.21	.14	.71	1.83	1.83	4.379
Rank ordered ass.	.021	.12	2.81	1.38	1.39	10.60	4.61	16.591**
MHC	.001	.56	.00	.17	2.41	1.39	1.39	5.18
Cascading ass.	.003	-.295	.00	.00	.00	1.39	1.39	2.774
MDS	.028	.54	2.03	.58	2.41	7.38	2.41	12.194**
Optimal ass.	.028	.04	1.62	1.64	1.39	5.99	5.99	13.368**

* p < .05 ** p < .01

Notes.

[a] **Assignment strategy abbreviations**: Random ass.: random order assignment; Rank ordered ass.: rank ordered jobs by validity; MHC: multidimension hierarchical classification; Cascading ass: cascading assignment; MDS: multidimensional selection; Optimal: optimal linear assignment

[b] **MD** denotes mean difference between operational & proportional cut score condition MPP standard scores (see Table 4) across 3 analysis sample conditions

[c] **CR** denotes critical ratio

Experiment 2

RESEARCH DESIGN

The second experiment of the present research study utilized the data and research methodology from Experiment 1 to measure assignment variables that result from the removal of Brogden's *g* from assignment variables computed as LSEs of the criterion variable. The same twenty samples of synthetic entities were used to remove Brogden's *g* and to simulate selection and assignment. The nine Project A jobs used in Phase 1 of Experiment 1 (Batch A jobs) and the 27 predictors comprising the ASVAB and

longitudinal data set (see Tables 2 and 3), as in Experiment 1, provided the data for the generation of alternative samples required for the model sampling methodology. "Population" intercorrelation and validity coefficient parameters obtained via the correction procedure reported for Experiment 1, were used to generate samples of predicted performance scores, which were least squares estimates of performance, for the purpose of partialling out Brogden's *g* from the LSE assignment variables separately in each of the 20 generated cross samples. As in Experiment 1, the analysis samples were used to calculate the regression weights use to define the LSEs. These weights were then applied to the test scores in the cross samples to provide the predicted performance scores used as the baseline assignment variables.

Experiment 2 involved an additional 6 conditions bearing on simulating selection and assignment of personnel to 9 Project A jobs. The variable of key interest in this experiment was the method of calculating the regression weights for the two types of assignment variables. Two alternative methods of obtaining these weights were contrasted under three different analysis sample sizes. Table 8 shows the two facets in Experiment 2 and the levels of each facet used to create the 6 simulation conditions. In contrast to the conditions in Experiment 1, no cut scores were applied to the sample test scores and only an optimal assignment (LP) procedure was used.

A baseline condition was provided by use of least squares estimate assignment variables as AVs. To provide the second level of this facet, Brogden's *g* was partialed out of the baseline assignment variables by removing a factor representing Brogden's *g* from the intercorrelation and validity coefficients used to compute the weights. The two assignment variable conditions were applied under three different analysis sample sizes: 405, 1620, and 3645. The two larger-sized analysis samples were already available from Experiment 1 and were reused making the differences in MPP provided by optimal assignment with, and without, the use of cut scores.

Four measures were used as indicators of the effect removal of Brogden's *g* has on characteristics highly correlated with classification efficiency. These measures are as follows: (1) any change in the mean predicted performance standard score obtained after optimal assignment during simulation (as for one of the assignment strategy options in Phase 1 of Experiment 1); (2) the amount of shrinkage in predictive validity (R); (3) the increase in the average intercorrelations among assignment variables (r), and; (4) increase in Brogden's (1955) estimate of classification efficiency (or MPP) which is calculated as the product of a function of the number of jobs, tabled by Brogden (1959), times the product of R and

a function of the intercorrelations of the assignment variables (i.e., r); this function of r is the square root of $(1 - r)$. That is, Brogden's estimate of MPP is equal to the square root of $(1 - r)$ times $f(m)$ times R. Brogden's tabled value for $f(m)$, for $m = 9$, is 1.49.

Table 8. *Experimental design (Experiment 2)*

Experiment 2 Variables	Levels
Assignment variable	1. Least squares estimate LSE) 2. Brogden's *g* removed
Analysis sample size	LSEs 1. N = 405 2. N = 1620 3. N = 3645

RESULTS

Estimates of MPP based on Brogden's (1959) assumptions are computed for LSEs as assignment variables (AVs) in making optimal assignments, and for a *g* factor distinct from the AVs that makes no contribution to classification efficiency. Intercorrelations among the AVs are assumed to be a function of this *g* factor.

The two levels of the AV type facet of Experiment 2, the traditional LSEs (i.e., baseline) variable vs. LSEs with Brogden's *g* partialed out, both meet only part of the assumptions of Brogden's model for estimating MPP. The first level AVs are true LSEs and thus include all aspects of psychometric *g* identifiable within the joint space, and the components of *g* found in the AVs do have an important impact on classification efficiency. The second level AVs are independent of Brogden's *g* as assumed in Brogden's model, but this *g* does not determine the intercorrelations of the AVs, as assumed in his model. In fact, less than half of the covariances among second level AVs are determined by the *g* variance that has been partialed out of the AVs.

Table 9 shows that both of the levels of the AV type facet provided the same values of MPP (see rows one and two). This relationship is so close that one can have little doubt that an analytical proof of the exact equivalence of these two levels for computing MPP may some day be produced. Such a proof would require proving the equivalence of some very complicated integrals and is not attempted in this

study. However, model sampling clearly indicates that Brogden's *g* has no effect on classification efficiency in the same sample used for making optimal assignments.

Another important finding of Experiment 2 can be seen from comparing MPPs computed in the absence of cut scores, as shown in Table 9, with MPPs computed in conjunction with the use of cut scores on AVs as a preliminary to the application of the LP algorithm, as shown in Table 4. For N = 1620, baseline LSEs used without cut scores provided an MPP of .7204, compared to .7497 with the better cut scores, while the same LSEs used with the N = 3645 analysis sample without cut scores provided MPPs of .7253 compared to .7762 with the better cut scores.

A cross sample loss in MPP of .02 in this repeated measures design is large enough to require an explanation, but is too surprising to entirely discount as a possible result of sampling error. An MPP computed in the analysis sample would necessarily provide a higher MPP for LSEs used as AVs in a LP algorithm without, as compared to with, cut scores. In the cross samples the application of optimal cut scores to the AVs, with the height of each cut score computed in the analysis samples, the interaction of cut scores and LP algorithm is not so predictable and the gain in MPP found when the operational cut scores are removed contrasted with the loss in MPP found when the cut scores proportional to the dual parameters are removed, strongly suggests that such cut scores optimally utilized could improve classification efficiency.

Table 9. *Experiment 2 simulation results across 20 replications*

	Analysis sample size		
Calculation of assignment variable	N = 405	N = 1620	N = 3645
Average MPP after optimal assignment			
Baseline: least squares estimates	.6671 (.1461)	.7204 (.1569)	.7253 (.1580)
Brogden's *g* removed	.6671 (.1461)	.7204 (.1569)	.7253 (.1580)
Average predictive validity across assignment variables			
Baseline: least squares estimates	.9303 (.1988)	.9717 (.2072)	.9760 (.2081)
Brogden's *g* removed	.4521 (.0983)	.4821 (.1052)	.4744 (.1041)
Average intercorrelation among assignment variables			
Baseline: least squares estimates	.5075 (.1081)	.5377 (.1153)	.5317 (.1141)
Brogden's *g* removed	.2524 (.0551)	.3091 (.0674)	.3293 (.0716)
Average Brogden MPP			
Baseline: least squares estimates	.4378 (.0963)	.4761 (.1044)	.4844 (.1061)
Brogden's *g* removed	.2608 (.0589)	.2889 (.0655)	.2821 (.0642)

Experiment 3

BACKGROUND

There is an extensive literature on weight stabilization in the context of selection or placement for a single job. However, in the case of multiple jobs, an assumption of independent applicant streams for each job is implied by the use of predictive validity as the unit of measurement in the determination of shrinkage. Many of the weight stabilization approaches, derived from research based on the objective of increasing predictive validity for a single job in a cross sample, rely on the reduction of dimensionality in the predictor space.

The reduction of dimensionality is clearly inconsistent with the objective of obtaining classification efficiency (and also, we contend, of multiple job selection efficiency in a system based on a common applicant pool). There is a need to explore the effect variables relevant to personnel

classification have on MPP in independent samples when weight stabilization methods consistent with DAT are used in the specification of assignment variables.

There appears to be a consensus among measurement scientists that the multiple regression weights yielding maximum validity in the back sample do not necessarily provide a larger validity in an independent "cross" sample as compared to a unit weighted composite. This consensus is especially strong when the size of the back sample is very small and the intercorrelations among predictors are high. Unfortunately, the large research literature on this topic, has almost entirely restricted itself to considering selection efficiency for jobs with independent applicant pools, as measured by predictive validity. The possibility arises that the stability problem might have entirely different findings when classification efficiency is the secondary goal of stability.

It has been suggested that the reason for weight instability is closely related to the increase in reliability that occurs as composite weights approach equality across the independent variables. Mosier proved (1943) that a weighted test composite has greater reliability when these intercorrelations are higher. It also seems logical that a test composite will show less shrinkage of predictive validity in cross samples as the standard error of regression weights decreases. It seems reasonable that a similar result would occur with respect to classification efficiency (MPP) as the sampling error of regression weights decrease. The greater sampling error found in regression weights when the correlations among dependent variables is higher seems to contradict the possibility that increasing composite reliability increases weight stability.

Some investigators expect greater stability from regression equations estimated from the largest PC factor as contrasted to the largest Horst differential factor, H_d, (that is almost always smaller than the largest PC factor). They are probably assuming that, as a general rule, the larger factor has proportionately less sampling error than does the smaller H_d factor. This assumption is disputed by Davis (1945) who challenges the myth that smaller factors necessarily contain a larger percentage of sampling error (as compared to larger factors). Statman's (1993) results show a considerable higher MPP for the composites based on the H_d factor. This finding supports the expectation that the removal of g from AVs, by some methods, might increase stability without reducing MPP. Guttman (1958) warns against assuming that reducing rank is always beneficial to the improvement of measurement.

Draper and Van Nostrand (1970, p. 436) challenge the use of ridge regression in his comments and opinion section. Project A investigators (McLaughlin, Rossmeissl, Wise, Brandt, & Wang, 1984)

make use of a ridge regression correction of regression weights in the preliminary stage of their analysis. The ridge regression correction has the effect of making the predictor inter-correlation coefficients more similar to each other, such as would happen if a matrix of constants were added (not subtracted as we do to extract *g*) from the multiple correlation matrix. This is the opposite direction as our subtraction of the effects of *g* from the multiple correlation matrix and would increase the similarity of regression weights for AVs. Marks (1977) proposes a method which increases the similarity among validity coefficients across predictors.

It is interesting to note that Rock, Linn, Evans, and Patrick (1970) and Schmidt (1971) came to opposite conclusions as to the usefulness of regression equations when the values of the regression weights are computed on small samples. Regardless of whose findings you choose to believe, there are still no data that bear on the desirability of adjusting regression weights for use as either selection or assignment variables in a selection or classification procedure involving multiple jobs that share a common applicant stream. Statman (1993) compares FLS composites using adjusted regression weights with composites based on weights computed within the common factor space included in the joint predictor-criterion space.

A method described by Curtis (1976) has potential value for classification because it substitutes validities for reliabilities and thus would provide distinct composites for each job family. Curtis provided cross validation results for a number of MOS in which the weights were obtained by the alternative method.

The literature on weight stabilization has almost entirely focused on the effect of methods that have the effect of reducing the range of least squares weights have on predictive validity. Most of the proposed methods appear intuitively to be counter productive with respect to classification efficiency.

Four of the traditional methods for weight stabilization are: (1) the elimination of negative weights; (2) the reduction in the number of predictors in a LSE through sequential test selection; and (3) the reduction in the dimensionality of test space by using only the larger of the principal components to define a reduced space within which to compute weights, and (4) the use of unit weights. Each of the four methods have been examined in DAT-oriented simulation experiments. No instance was found of any other type of AV showing an increase in MPP over the use of FLS composites as AVs.

It has been fairly well established that the proportional size of the first principal component in the joint predictor-criterion (JP-C) space, as compared to total factor contributions, is usually larger than the proportional size of the first principal component in test space. The following two principles also have been generally accepted:

1. The first principal component in the JP-C space will typically provide 80 percent, or more, of the total factor contributions.
2. The combination of principal components other than the first will almost certainly provide more classification efficiency than the first principal component.

Obtaining of best weights for the purpose of maximization of the reliability of a test composite is described by Mosier (1943) and again, more generally, by Peel, E. A. (1948). (Also see Gulliksen, 1950, pp. 346-351.) Note that in Peel's solution the N by m matrix [R] is the intercorrelations among our n AVs with ones in the diagonals while [C] is the same matrix with reliability coefficients in the diagonals. Curtis substituted a matrix [J] for [C] where [J] has validity coefficients, instead of reliabilities, in the diagonals. Obviously there is a separate matrix [J] for each of our 9 MOS.

METHOD

Research Design

The third experiment in this project utilized the data and research methodology from Experiment 1 to measure the stability of two alternative computations of assignment variables. The same twenty samples of synthetic entities were used to simulate selection and assignment under alternative conditions. The nine Project A jobs used in Phase 1 of Experiment 1 (Batch A jobs) and the 27 predictors comprising the ASVAB and longitudinal data set (see Tables 2 and 3), as in Experiment 1, provided the data for the generation of alternative samples required for the model sampling methodology. "Population" intercorrelation and validity coefficient parameters obtained via the correction procedure reported for Experiment 1, were used to generate samples of predicted performance scores, which were least squares estimates of performance, for the purposes of evaluation across conditions. Likewise, three different sized analysis samples, as in Experiment 2, were used to calculate assignment regression weights for predicted performance scores.

Experiment 2 involved an additional 12 conditions simulating selection and assignment to Project A jobs. The variable of key interest in this experiment was the method of calculating the

regression weights for the full least squares estimate assignment variables. Four alternative methods of obtaining these weights were contrasted under three different analysis sample sizes. Table 8 shows the two variables in Experiment 2 and the levels of each variable used to create the 12 simulation conditions. In contrast to the conditions in Experiment 1, no cut scores were applied to the sample test scores and only an optimal assignment procedure was used.

A baseline condition was provided by least squares estimate assignment variables. Two alternative stabilization methods were created as modifications of these least squares estimates using different ways of creating factor-based regression weights: (1) a model for maximizing reliability of the test composite comprising the full set of predictors was maximized using validity coefficients for distinct jobs in place of reliabilities in order to provide distinct composites across job families; and (2) the same method of increasing test composite reliability as the previous condition was approximated using a principal components factor solution in the predictor space.

The two assignment variable conditions were applied under three different analysis sample sizes: 405, 1620, and 3645. The two larger-sized analysis samples were already available from Experiment 1.

Each condition was evaluated for its usefulness in increasing the stability of the regression estimate and ultimately improving classification efficiency. Three measures were investigated as initial indicators of improved usefulness of AVs based on experimental methods for modifying the regression estimates: (1) change in predictive validity (R) across conditions; (2) change in the average intercorrelations among assignment variables (r); and increase in Brogden's (1959) estimate of classification efficiency (or BMPP) which is calculated as a function of R, r, and $f(m)$. If these changes had been in the direction of a decrease in r and an increase in R we would have also expected an increase in BMPP and would have also conducted the simulations to measure the change in the increase in the mean predicted performance standard score obtained after optimal assignment during simulation (i.e., SMPP).

RESULTS

No experimental method for improving the reliability of composites or for increasing the stability of AV regression weights produced AV intercorrelations (i.e., r) below a range of from .90 to .97, while increasing the validity coefficients in cross samples only moderately. Thus the values for

BMPP was decreased to being less than 10 percent of the value provided by the LSE AVs. It was not considered worthwhile to conduct simulations using these modified AVs.

Discussion and Conclusions

The present study is a basic research effort hopefully will both stimulate other more applied research and add to the known differential assignment theory (DAT) principles (Johnson, and Zeidner, 1995). There are a number of important DAT related issues that bear on the findings of this study. The most important of these are as follows: (1) there appears to be worthwhile gains in MPP that can be attributed to the use of cut scores proportionate to dual parameters instead of to judged "difficulty" of the MOS; (2) the results show a continuous progression of MPP gains that show classification efficiency approaching that of line-by-line optimization models as multiple job selection models are made more and more similar to the dual LP algorithm; (3) AV sets that have had Brogden's *g* removed in a back sample situation for the purpose of obtaining predictor weights for AVs yield essentially the same assignments as do the traditional AVs; (4) there is a complete lack of promise for improving classification efficiency provided by weight stability methods that rely on reducing the dimensionality of the AVs; (5) there appears to be a relatively small loss of MPPs attributable to using minimum cut scores as a preliminary step to the use of an LP algorithm--regardless of which of the two kinds of cut scores are utilized; and (6) the small but significant gain provided by using cut scores proportional to the dual parameters, as compared to the operational cut scores, confirm the hypothesis of superiority--especially if the cut scores are applied to adjusted scores instead of directly to AV scores.

The DAT literature provides a number of ways to increase MPP from a classification process in second best situations where policy forbids the use of the ideal approach. For example, the maximum MPP is obtained using full least squares estimates based on all predictors but DAT provides a way to maximize MPP when composites are limited to a smaller number (e.g., 3 or 4), or by special approaches when all weights must be positive. Similarly, there are ways to maximize MPP by optimally clustering jobs into families for a prescribed lower number of job families than the truly optimal number that is obtained by making every MOS with adequate validity data into a separate family.

Information on the cost in reduced MPP that can be expected to result from establishing non-optimal policy should be closely considered. In some situations trade-offs between maximizing MPP and other benefits such as: (1) reducing school or on-the job failures and behavioral problems; (2)

increasing the retention rate; or (3) the provision of leaders for critical missions, are also appropriately considered.

UNIDIMENSIONAL ASSIGNMENT VARIABLES

The superiority of univariate factor scores based on a Horst differential validity factor (H_d) as compared to those based on a principal component factor, both multiplied by the validity coefficients separately corresponding to the job families corresponding to each AV, is well documented as a source of MPP gains. Statman (1993), using Project A concurrent data, reported MPP standard scores of .737 for H_d as compared to .315 for a PC based factor score that many would say is a measure of *g*.

Statman's AVs were computed within a joint predictor-criterion space with dimensionality ranging from one to eight. The superiority of an eight dimensional space over a unidimensional measure for creating AVs was indicated by a 100 percent gain in MP for the AVs from the larger space. Our Experiment 1, Phase 2 results provide similar conclusions regarding the importance of multidimensionality in AVs and the lower MPPs provided by the unidimensional AVs approximating *g*, as compared to H_d.

The assignment variables employed in this study that are not measures of single factors are measured in the total joint predictor-criterion space, rather than, as in Statman's study, by a sub-set of this joint space spanned by 10, 8, 4, 2 or 1 factors. The superiority of the AVs based on the total joint space incorporates all of the ASVAB and Project A experimental tests in the AV best weighted composites (i.e., LSEs), is greater than is shown by Statman's study. This study focuses on the potential gains in MPP that are obtainable from an optimal selection of predictors in the context of both single and multi-factor models.

The most credible analytical measure of MPP as an alternative to the estimating MPP by means of a simulation of an optimal assignment process is provided by Brogden's MPP model (Brogden, 1959). Earlier we referred to this analytical estimate as, BMPP, and the simulation estimate as SMPP. BMPP is defined as $f(m)\ R\ (1\text{-}r)^{½}$, where R is the mean validity and r the mean intercorrelation among the AVs. A table with values for $f(m)$ up to $m = 15$ are provided by Brogden. Using the LSE type AVs obtained from the analysis sample with N = 3645, and the 20 replication cross samples, it was found that BMPP correlated .9592 with SMPP.

When BMPP is computed using AVs that have Brogden's g partialed out, the resulting values are underestimates of SMPP, in contrast to the estimates obtained from the use of r and R directly computed from the AVs that provide much closer approximations.

A set of optimal AVs associated with each of the above four single factors can be obtained by computing each individuals LSE score for each of the four factors, and multiplying the factor scores by the validity of each factor for each MOS criterion. Thus, each individual (or entity) has a separate vector of AV scores associated with each factor and a standard primal LP algorithm can be used to make optimal assignments in the same way the LSEs are utilized as AVs. However, except for the use of the preliminary minimum cut scores, the first three levels of the assignment facet could be closely approximated by an analytical algorithm for the single factor AVs, but not when LSEs are utilized as the AVs. The hierarchical classification model (HC) is essentially an analytical counterpart of the second level assignment algorithm in the absence of cut scores.

It was already known from the results of Statman (1993) that the use of either H_d or H_a provided less than half as high a value of MPP as did the next few factors. Also, it was known that the percentage increase provided by a single H_d factor was substantially larger than that provided by a single H_a factor. The results of this study provide from 9 to 10 times as high a value of MPP for AVs consisting of LSEs as compared to sets consisting of a single factor score times MOS validities.

Assuming that a single factor is to be utilized for selection and not for assignment to jobs, it is clear that either H_a or g factors have more validity and consequently provide more selection efficiency than would use of an H_d factor. When classification efficiency in the assignment process is the objective, the H_d factor is half again as efficient (in terms of MPP) as either H_a or g. The best set of univariate AVs for use in simultaneous selection and classification presents a more difficult choice and would depend on the selection ratio.

MULTIPLE JOB SELECTION STRATEGIES

Assignment of new recruits to alternative job families is simulated in this study using either 9 or 14 MOS as surrogates for families. All multiple job selection strategies evaluated in this study utilize AFQT scores to reject applicants from an applicant sample drawn from the youth population. This is the first stage, a selection stage, of a two stage process in which the second stage involves the distribution of new soldiers to job families. It would have been feasible to make use of the multidimensional screening

model (MDS) to maximize MPP while simultaneously selecting into the Army and assigning to jobs (Johnson and Zeidner, 1990). Whetzel (1991) demonstrated the superiority of MDS over the use of the traditional two stage model as used in this study. However, the authors believe the advantages of being able to use separate, more appropriate, criteria for each of the first and second stages easily makes up the loss in total MPP resulting from using the less efficient two, instead of one, stage model.

The most primitive of the selection strategies investigated in this study (level 1 of the strategy facet) considered both jobs and persons in random order, rejecting or accepting each person-job match, depending upon whether the applicable cut score was failed or met by the person under consideration. Each succeeding level was more selective regarding the order of one or both of persons or jobs in the making of a match. Filling the more valid jobs with the more able applicants is one way to provide a gain in MPP. This concept is the basis for the Hierarchical classification (HC) model that can function using either AVs based on a single predictor multiplied by different validity coefficients or AVs consisting of separate predictors for each job (Johnson and Zeidner, 1990).

All five of the most primitive strategies, the first through the fifth levels, are line-by-line algorithms in which assignments can be made for one individual at a time, in contrast to making the assignment decisions in a batch mode. Such line by line algorithms are referred to by the Air Force as "sequential" algorithms. Ward (1958) proposed one such algorithm. The sixth level of this facet that uses the primal LP algorithm requires assignments to be made in a batch mode (a separate batch for each replication).

The fifth level differs from the fourth level in that the dual parameters applied in the algorithms of both levels are applied to different variables. This application is directly applied to the AVs in the fourth level as contrasted to the "adjusted AVs" in the fifth level. Thus we see that the direction of increased complexity, going from level one to level five, leads to a convergence on a more efficient algorithm -- to a dual, line by line, LP algorithm. Level six consists of a primal LP algorithm that requires the simultaneous assignment of a moderately large sized sample of recruits.

The gains in MPP going from level 1 through level 5 are substantial when the cut scores proportional to dual parameters are utilized, but are also evident when the operational cut scores are used. Variants of these simpler algorithms were used in conjunction with IBM card sorters in the basement of the Pentagon during what amounted to a pre-computer age for the assignment process.

MINIMUM CUT SCORES

The low average height (low range) of the cut scores and their containment within a small range is not due to low judged difficulties of Army MOS. The minimum cut scores for the Army's operational test composites (AAs) reflect a long existing policy which emphasizes the desirability of meeting the assignment preferences of as many recruits as possible. To this end, the minimum cut scores are all placed in such a low range that most recruits can meet all of the cut scores and only rarely is a preferred MOS placed out of reach of a recruit during the assignment process.

This study uses LSEs instead of AAs as AVs for the simulated assignment process. Since these LSEs are stated with respect to predictors that have been placed in statistical standard score form, the operational cut score for each MOS is appropriately converted by subtracting 100 and dividing by 20.

Although the range of operational cut scores fall mostly between 90 (statistical standard score = -.50) and 110 (statistical standard score = + .50) with the total range falling between .85 (statistical standard score = -.75) and 120 (statistical standard score = +1.0). Approximately seventy percent of all Army minimum cut scores fall at or below 100 (statistical standard score = 0). Although it would appear that very little impact can be had by the operational use of such a low set of cut scores, considerable attention is given to the order each MOS receives within this limited range. This order reflects the order of perceived "difficulty" among the MOS. "Difficulty" among the MOS is estimated in terms of a number of factors including: (1) failure rate in Army Schools; (2) frequency of unsatisfactory on-the job performance; (3) by both safety experiences and by the damage to materiel and personnel that failure on the job might cause; and (4) the perceived importance of performance on the job to the unit's mission. The predicted need for soldiers experienced in a given MOS to be transferred to a more difficult MOS may be considered in setting cut scores. Obviously the decisions behind setting the operational cut scores are complex and bear on potentially important outcomes.

The operational minimum cut scores for the set of 14 MOS of this study are contained within a standard score range of .75. The policy objectives relating to the protection against very poor performers being assigned to certain MOS would probably be better met by increasing this range by a factor of as much as 2 or 3, but at a cost of reducing the percent of preferences met. However, little is known concerning the gain in performance that might be obtainable from increasing the range of the operational cut scores without also making a change in the order of the heights across MOS.

While it is clear that increasing the range of cut scores would decrease the number of preferences met, it is equally clear that the present judgmentally determined order of operational cut scores are not optimal for providing an increase in MPP. Prior to this study, there was no evidence that a different set of cut score heights would reject a set of applicants that would provide higher MPPs for the remaining soldiers -- as compared to the existing operational cut scores that have been selected as estimates of MOS "difficulty". Army policy has not placed much importance on keeping MPP high as a product of cut scores. Instead Army policy has led to a definition of an optimal set of cut score heights as ones which will provide protection from the more extreme of the poor performers while rejecting no more than a permissible number of applicants during the application of cut scores.

If minimum cut scores were to be applied to adjusted scores defined as AVs scaled within each person by subtracting each person's largest AV score from each of his AV scores, some variables proportional to the dual parameters (i.e., to the allocation constants) would be the most promising candidates for increasing MPP as the consequence of using minimum cut scores. The use of such ideal cut scores in a multiple job selection process would approach the MPP obtainable from the use of an LP assignment model. It is less clear that a cut score proportional to dual parameters could be as useful when directly applied to the AV scores. Unfortunately, cut scores applied to adjusted scores would not provide a minimum score across individuals bestowing protection of the units from the assignment of poor performers to the more "difficult" MOS.

The feasibility of increasing potential MPP by optimizing the pattern of cut scores is investigated in this study. Intuitively, there should be an optimal mean height, variance, and pattern of cut scores for a given set of job families (or surrogates). The investigators also hypothesize that up to some critical point an increase in overall mean cut score and in variance of cut scores across the job families increases MPP as compared to the operational cut scores. Intuitively, there should be a relationship between the pattern and mean height of the best set of cut scores and dual parameters that can effect a maximization of the mean AVs when a dual LP algorithm is applied to optimize assignments.

When an LP algorithm is used to make optimal assignments in an assignment sample, the objective function consists of the mean AV score, obtained by averaging the AV scores corresponding to the jobs to which each entity is assigned. This maximization of mean AV scores is made while meeting quota constraints (and any other constraints). Sorenson (1965) using his partially biased research design in which the same weights were used for AVs and EVs, found across a number of simulation

experiments that the size of the cross samples (N > 200) in which the LP based assignments were made had little effect on magnitude of MPP obtained from a simulation. This was a very convenient finding since an effective design for a research experiment requires 20 or more research replications and the cost of applying an LP algorithm goes up rapidly as the cross sample N increases. Since Sorenson obtained these findings, DAT based studies, with the notable exception of Nord and Schmitz (1987), have used cross samples with sample sizes lying between 300 and 500.

Determining the dual parameters on a large independent sample for use in a dual LP algorithm can provide a larger MPP, although a smaller objective function, than can be obtained by computing either a primal LP on a relatively small assignment (cross) sample or, conversely, determining the dual parameters on that small assignment sample. While the objective function of an LP algorithm will obviously be larger when computed in the same sample as the objective function is maximized, the MPP based on EVs will not necessarily be, since this source of bias has been removed.

The findings of this study regarding an increase in MPP from the preliminary use of cut scores proportional to the dual parameters (obtained from an independent sample) supports the hypothesis described in the paragraph immediately above. A demonstration by Granda and McMullen (1974) in which dual parameters were applied and followed up by the use of a primal LP algorithm, provided surprisingly high MPP values. These investigators made all assignments, except for individuals with ties and near ties before applying the primal LP algorithm.

In summary, shrinkage from a biased to an unbiased estimate of the population MPP can be expected to occur whenever there is any opportunity to capitalize on sampling error. Some of these opportunities are more familiar than others to the extent estimation of dual parameters is a non-trivial source of such a bias, the population MPP can be increased by at least partially estimating optimal assignments using parameters obtained from an independent back sample, especially when the back sample is larger than the assignment sample. This conclusion has implications for the utility of using a line-by-line optimal assignment (dual LP) algorithm, making such an algorithm even more advantageous as compared to the more conventional LP algorithm. It appears likely that an increased magnitude of MPP may be provided, in addition to the convenience of making assignments for each recruit in one-on-one interactions between recruits and recruiters or counselors.

FOUR KINDS OF UNIDIMENSIONAL FACTOR SCORES

Four different unidimensional factors are demonstrated and explored in this study. The Brogden *g* factor is identified in the joint predictor-criterion space as the largest factor which has equal validities with each of the 9 or 14 MOS criterion variables. Any further increase in the factor contribution of this factor would result in a negative eigen value in the joint space indicating that this space can no longer be considered as Euclidian.

The largest principal axis factor computed from a correlation matrix of 29 predictors (i.e., defining the test space) provides a factor that is commonly referred to as psychometric *g*. The latter, referred to as *g* in Table 5, is much larger than Brogden's *g* and has a considerable range of validities against the separate MOS. A third factor, referred to as H_a in Table 5, is the largest principal axis factor obtainable from the covariances among the AVs (computed as LSEs separately for each MOS) of the *m* (9 or 14) criterion variables. The covariances among these LSEs define the joint predictor-criterion space. This *m* by *m* covariance matrix has the squared validities of the LSEs in its diagonals. The resulting H_a factors reproduce the covariances in this joint space as contrasted to the reproduction of the correlations among the 29 tests provided by the *g* factor.

The fourth factor does not measure another concept of *g*, as do the first three, but instead maximizes differential validity using a single factor. This is the largest of Horst's differential validity factors that sequentially maximize Horst's index, H_d.

A set of optimal AVs associated with each of the above four single factors can be obtained by computing each individual's LSE score for each of the four factors, and multiplying the factor scores by the validity of each factor for each MOS criterion. Thus, each individual (or entity) has a separate vector of AV scores associated with each factor and a standard primal LP algorithm can be used to make optimal assignments in the same way the LSEs are utilized as AVs. However, except for the use of the preliminary minimum cut scores, the first three levels of the assignment facet could be closely approximated by an analytical algorithm for the single factor AVs, but not when LSEs are utilized as the AVs. The hierarchical classification model (HC) is essentially an analytical counter part of the second level assignment algorithm in the absence of cut scores.

It was already known from the results of Statman (1993) that the use of either H_d or H_a provided less than half as high a value of MPP as did the aggregate of the next few factors. Also, it was known

that the percentage increase provided by a single H_d factor was substantially larger than that provided by a single H_a factor. The results of this study provide from 9 to 10 times as high a value of MPP for AVs consisting of LSEs as compared to sets consisting of a single factor score times MOS validities.

Assuming that a single factor is to be utilized for either or both selection and assignment to jobs, it is clear that either H_a or g factors have the most validity and provide more selection efficiency than would an H_d factor. When classification efficiency in the assignment process is the objective, the H_d factor is half again as efficient (in terms of MPP) as either H_a or g. The choice of the best set of univariate AVs for use in simultaneous selection and classification presents a more difficult decision.

ASSIGNMENT VARIABLES ORTHOGONAL TO BROGDEN'S g

It would be convenient to use the same set of AVs in both a computer based optimal assignment system and for vocational counseling procedures. Most counselor's would agree that a set of AVs with smaller intercorrelations would add to the acceptability of the AVs for their purposes. Unfortunately, the LSEs of the criterion used as AVs to maximize MPP obtainable from an optimal assignment process typically results in high intercorrelations among such AVs. The use of factor scores derived from up to 10 factors that have been rotated to simple structure has been demonstrated by Statman (1993). This method, which calls for computing AVs as LSEs of factor scores instead of criterion scores, greatly reduced the intercorrelations among the AVs (based on 29 Project A tests) at a cost of only a small reduction in MPP. It would be desirable to have a more general method that can provide a similar reduction in intercorrelations with a minimum reduction in MPP. Hopefully such a method could be applied to larger numbers of job families and/or fewer predictors without harm to its desirable characteristics.

Brogden's (1959) MPP model has assumptions which imply the applicability of Spearman's two-factor model. Brogden's model includes the description of covariances among m AVs and m criterion variables using $m + 1$ factors, including a general factor and m unique factors. This factor solution obtained in the joint predictor-criterion space corresponds to Spearman's (1927) two-factor theory. The covariances among the AVs are entirely explained by the single general factor. The length of each AV vector (corresponding to each AV validity) are determined by the square root of the sums of squares of the loadings of each AV on the g factor and on the factor unique to that AV. The assumptions implied by this model permitted Brogden to analytically estimate MPP without the necessity of solving intractable (when $m > 3$) multivariate integrals. Brogden's estimate of MPP, referred to here as BMPP, is invariate

with respect to the size of AV loadings on his *g* factor. Thus the extraction of *g* from the AVs would have no effect on the magnitude of BMPP when Brogden's assumptions are met. Using the formula BMPP = $f(m)\ R\ (1-r)^{1/2}$, the extraction of Brogden's *g* reduces *r* (the intercorrelations among AVs) while also reducing *R* (average validity of AVs). However, the certainty that AVs equal to LSEs computed on empirical data do not equal Brogden's assumptions make it necessary to seek other evidence than is obtainable from Brogden's model regarding whether MPP is invariant to Brogden's *g*.

MPP obtained by use of an unbiased simulation of assignments to multiple jobs is referred to as SMPP. The difference between SMPP computed before and after the extraction of Brogden's *g* from the AVs provides evidence bearing on presence or absence of invariance with respect to Brogden's *g*. Such evidence is provided by the results of Experiment 2. SMPP remained the same whether the unaltered LSEs or LSEs with the effects of *g* removed were utilized as AVs. However, this invariance is with respect to a situation in which Brogden *g* is extracted from each AV and assignments optimized in the same sample. If *g* had been extracted in the process of computing the weights for the AVs in the analysis samples, the independence of the sampling error in the analysis and cross samples would have introduced differences between the two types of AVs.

A second question of interest was whether the computation of BMPP using *R* and *r* computed using the AVs orthogonal to *g* would provide a better estimate of SMPP than would be provided by BMPP based on the *R* and *r* computed from the unaltered AVs. The results of this study show that when unaltered AVs were used to compute average values of *R* and *r* ,and to then to compute BMPP, provided very close estimates of SMPP. In contrast, when AVs orthogonal to *g* were used to compute average values of *R* and *r* and to then compute BMPP, considerable underestimates of SMPP were found. Equally large over estimate of SMPP were provided from the separate estimates of BMPP in each cross sample provided by the use of unaltered AVs.

The intercorrelations among sample values of *R*, *r*, BMPP, and SMPP, computed across the 20 cross samples provided limited information. The 20 cross samples differ from each other in terms of sampling fluctuations and do not provide insight into intercorrelations among key variables if each sample had been drawn from a different population having a different true value of SMPP. However, BMPP and SMPP correlate in the upper nineties across the 20 cross samples and would necessarily have an even higher correlation for samples drawn from different populations with different values of SMPP.

Thus BMPP would have value as an estimate of SMPP in situations where the latter was not available, but the authors would not propose the use of BMPP as a substitute for SMPP in most research situations.

CONCLUSIONS

The primary conclusions of this study can be formed as answers to the following questions:

1. Will use of minimum cut scores always reduce MPP obtained from an optimal assignment (LP) algorithm?
2. Can MPP be increased by using a line-by-line LP algorithm, instead of using a traditional primal LP algorithm?
3. Can LSEs be modified so as to substantially reduce the intercorrelation (r) among AVs, without reducing MPP?
4. Can hierarchical classification (HC) models using an unidimensional set of AVs (e.g., AVs equal to g multiplied by the validity of g for each job) provide a realistic alternative to the use of LSEs as AVs?
5. Could minimum cut scores determined on the appropriate AVs be optimized (in pattern of heights) and mean heights so as to maximize MPP, much as the weights of AVs can be optimized?
6. Can weight stabilization methods increase the magnitude of MPP in cross samples, or is the weight stabilization concept workable only with respect to the goal of increasing MPP as a result of selection?

The answer to question 1 is found in the finding in a prior study (Nord & Schmitz, 1991) that the use of operational cut scores reduce MPP by a small, almost trivial, amount. The finding in the present study that a preliminary application of cut scores proportional to the dual parameters can actually increase MPP makes it reasonable to infer that the direction of change in MPP imposed by the preliminary use of cut scores is a function of whether the height and pattern of cut scores are determined to maximize selection or classification efficiency.

The answer to question 2, one bearing on the potential efficiency of line-by-line assignment algorithms, is yes, for the same reasons as are given in response to question 6. The authors would also answer question 3 with a yes, but only when Brogden's g is removed from the AVs in the same sample where the optimal assignments are determined. Further research is required to determine if similar desirable results are obtained when the assignments are determined in a sample independent of the sample where Brogden's g is removed. Thus, question 3 has not been completely answered. The removal of the effects of Brogden's g from the weights for the AVs in the analysis sample suggests the possibility that a kind of weight stability very useful to classification efficiency might be added while

possibly reducing the reduction in the average intercorrelation among the modified AVs. Additional research on this question would be profitable.

The answer to question 4 is a resounding no. The MPP provided by LSEs is as much as 9 or 10 times as large as that provided by AVs defined as a function of *g* and the MOS validities.

Question 5 cannot be answered separately regarding pattern and mean heights. One level of the cut score facet has both a higher mean height and a pattern of heights hypothesized as superior for providing classification superiority. The higher MPPs for this level could be due to either variable, or both.

The answer to question 6 derives from the DAT principle that increasing mean validity (R) in a cross sample at the cost of increasing the intercorrelations (r) will almost always result in a decrease in MPP. The traditional methods for weight stabilization, as used by the authors in this and a previous study, have resulted in such an increase in r as to insure, despite a significant increase in R, a drastic reduction in MPP.

RECOMMENDATIONS

It is recommended as a moderately long range objective that separate predictor variables for the application of cut scores of the two types of cut scores be utilized. Predictor composites with high selection efficiency should be utilized in conjunction with "difficulty" cut scores to best meet the objectives of the existing minimum cut scores. Cut scores proportional to dual parameters, but of optimal mean height, should be applied to LSE type AV scores.

It is further recommended as a longer range objective that further research be conducted on the removal of Brogden's *g* from LSE type AVs with the objective of providing AVs with lower intercorrelations for use in counseling new recruits.

References

Alley, W. E., & Darby, M. M. (1995). Estimating the benefits of personnel selection and classification: An extension of the Brogden table. *Educational and Psychological Measurement, 55*, 938-958.

Brogden, H. E. (1946). An approach to the problem of differential prediction. *Psychometrika, 11*, 139-54.

Brogden, H. E. (1954). A simple proof of a personnel classification theorem. *Psychometrika, 19*, 205-208.

Brogden, H. E. (1959). Efficiency of classification as a function of number of jobs, percent rejected, and the validity and intercorrelation of job performance estimates. *Educational and Psychological Measurement, 19*, 181-90.

Curtis, E. W. (Dec., 1976). *Factor regression analysis: A new method for weighting predictors* (NPRDC TR 77-12). San Diego, CA: Navy Personnel Research and Development Center.

Davis, F. B. (1945). The reliability of component scores. *Psychometrika, 10*, 57-60.

Draper, N. R., & Van Nostrand, R. C. (1970). Ridge regression and James Stein estimation: Review and comments. *Technometrics, 21*, (4), 451-466.

Dwyer, P. S. (1954). Solution of the personnel classification problem with the method of optimal regions. *Psychometrika, 19*, 205-208.

Granda, T. M., & McMullen, R. L. (1974). *A heuristic approach to a large personnel assignment problem.* Paper presented at the Fall ORSA/TIMS National Convention.

Guttman, L. (1958). To what extent can communalities reduce rank? *Psychometrika, 23*, 297-308.

Fisher, R. A. (1948). Combining independent tests of significance. *American Statistician, 2*(5):30.

Gass, S. I. (1975). *Linear programming: Methods and applications* (3rd ed.). New York: McGraw-Hill.

Gulliksen, H. (1950). *Theory of mental tests.* New York: Wiley.

Johnson, C. D., & Zeidner, J. (1990). *Classification utility: Measuring and improving benefits in matching personnel to jobs* (IDA Paper P-2240). Alexandria, VA: Institute for Defense Analyses.

Johnson, C. D., & Zeidner, J. (1991). *The economic benefits of predicting job performance. Vol. 2: Classification efficiency.* New York: Praeger.

Johnson, C. D., & Zeidner, J. (July, 1995). *Differential assignment theory sourcebook* (ARI Research Note 95-43). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Johnson, C. D., Zeidner, J., & Leaman, J. (1992). *Improving classification efficiency by restructuring Army job families* (Technical Report 947). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Johnson, C. D., Zeidner, J., & Scholarios, D. (1990). *Improving the classification efficiency of the Armed Services Vocational Aptitude Battery through the use of alternative test selection indices* (IDA Paper P-2427). Alexandria, VA: Institute for Defense Analyses.

Johnson, C.D., Zeidner, J., & Scholarios D. (July, 1995). *Developing new test selection and weight stabilization techniques for designing classification efficient composites* (ARI Research Note 95-42). Alexandria, VA: U.S. Army Research Institute.

Marks, M. R. (1966). *Two kinds of regression weights which are better than betas in crossed samples.* Paper presented at Division 5 meeting of the American Psychological Association, New York.

McLaughlin, D. H., Rossmeissl, P. G., Wise, L. L., Brandt, D. A., & Wang, M. M. (October, 1984). *Validation of current and alternative ASVAB area composites, based on training and SQT informational of FY81 and FY82 enlisted accessions* (Technical Report 651). Alexandria, VA: U.S. Army Research Institute.

Mosier, C. I. (1943). On the reliability of a weighted composite. *Psychometrika, 8*, 161-168.

Niehl, E. & Sorenson, R. C. (January, 1968). *Simpo-I entity model for determining the qualitative impact on personnel policies* (Technical Research Note 193). Washington, DC: Statistical Research and Analysis Division, U.S. Army Behavior and Systems Research Laboratory. (AD 831 268).

Nord, R., & Schmitz, E. (1991). Estimating performance and utility effects of alternative selection and classification policies. In J. Zeidner & C. D. Johnson (Eds.), *The economic benefits of predicting job performance, Vol. 3. Estimating the gains of alternative policies* (pp. 73-137). New York: Praeger.

Olson, P. T. Sorenson, R. C., Hayman, K. W., Witt, J. M. & Abbe, E. N. (1969). *Summary of Simpo-I model development* (Technical Research Report 1157). Washington D.C.: U.S. Army Behavioral Science Research Laboratory.

Peel, E. A. (1948). Prediction of a complex criterion and battery reliability. *British Journal of Psychology, 1*, 84-94.

Rock, D. A., Linn, R. L., Evans, F. R., & Patrick, C. A. (1970). Comparison of predictor selection techniques using Monte Carlo methods. *Educational and Psychological Measurement, 30*, 873-884.

Schmidt, F. L. (1971). The relative efficiency of regression and simple unit predictor weights in applied differential psychology (Doctoral dissertation, Purdue University, 1970). *Dissertation Abstracts International, 31*, 6969B-6970B.

Scholarios D., Johnson C. D., & Zeidner, J. (1994). Selecting predictors for maximizing the classification efficiency of a battery. *Journal of Applied Psychology*, 412-424.

Sorenson, R. C. (1965). *Optimal allocation of enlisted men - Full regression equations vs. aptitude area scores* (Technical Research Note 163). Washington, D.C.: U.S. Army Personnel Research Office.

Spearman, C. (1927). *The abilities of man.* New York: MacMillan.

Statman, M. A. (1992, August). *Developing optimal predictor equations for differential job assignment and vocational counseling.* Paper presented at the annual convention of the American Psychological Association, Washington, DC.

Statman, M. A. (1993). *Improving the effectiveness of employment testing through classification: Alternative methods of developing test composites for optimal job assignment and vocational counseling.* Unpublished doctoral dissertation, The George Washington University.

Ward, J. H. (1958). The counseling assignment problem. *Psychometrika*, 23, 55-65.

Whetzel, D. L. (1991). *Multidimensional screening: Comparison of a single-stage personnel selection/classification process with alternative strategies.* Unpublished doctoral dissertation, The George Washington University.